

Contents lists available at [SciVerse ScienceDirect](http://SciVerse.ScienceDirect.com)

Applied Mathematics Letters

journal homepage: www.elsevier.com/locate/aml

An iterative least squares estimation algorithm for controlled moving average systems based on matrix decomposition[☆]

Huiyi Hu^a, Feng Ding^{a,b,*}

^a Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, PR China

^b Control Science and Engineering Research Center, Jiangnan University, Wuxi 214122, PR China

ARTICLE INFO

Article history:

Received 30 April 2012

Received in revised form 25 June 2012

Accepted 26 June 2012

Keywords:

Least squares

Iterative method

Parameter estimation

Controlled moving average model

ABSTRACT

An iterative least squares parameter estimation algorithm is developed for controlled moving average systems based on matrix decomposition. The proposed algorithm avoids repeatedly computing the inverse of the data product moment matrix with large sizes at each iteration and has a high computational efficiency. A numerical example indicates that the proposed algorithm is effective.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The least squares methods are important for linear function fitting, nonlinear function fitting [1], system identification [2–4], state estimation and filtering [5–7], and adaptive control [8–12]. Recently, many order identification and parameter estimation methods have been reported for linear or nonlinear systems [13,14], e.g., the least squares methods [15,16], the stochastic gradient methods [17–19], the multi-innovation identification methods [20–23] and the iterative methods [24–26]. The iterative methods are very important for finding the solutions of matrix equations or the roots of nonlinear equations [27,28], e.g., the Jacobi iteration and Gauss–Seidel iteration methods [29,30]. Ding et al. presented a least squares based iterative identification method and a gradient based iterative identification method for OEMA models [31]; Liu et al. developed a least squares based iterative identification method for Box–Jenkins models [32]; Bao et al. proposed a least squares based iterative parameter estimation algorithm for multivariable controlled ARMA systems [33]; Xie et al. gave a gradient based iterative algorithm and a least squares based iterative algorithm for matrix equations $AXB + CX^T D = F$ [34,35]; Zhang et al. derived a hierarchical gradient based iterative estimation algorithm for multivariable output error moving average systems [36]; Wang presented a least squares based recursive and iterative estimation algorithm for OEMA systems [16]. Ding et al. studied the hierarchical least squares identification for linear single-input single-output systems with dual-rate sampled data [37] and the bias compensation based parameter estimation for output error moving average systems [38].

On the basis of the decomposition based iterative estimation algorithm for autoregressive moving average (ARMA) models given in [39], this work derives an iterative least squares identification algorithm for controlled moving average (CMA) systems or finite impulse response moving average (FIR-MA) systems based on matrix decomposition with finite data length.

[☆] This work was supported in part by the National Natural Science Foundation of China (No. 60973043), the Natural Science Foundation of Jiangsu Province (China) and the 111 Project (B12018).

* Corresponding author at: Control Science and Engineering Research Center, Jiangnan University, Wuxi 214122, PR China.

E-mail addresses: hhygut@sina.com (H. Hu), fding@jiangnan.edu.cn (F. Ding).

The rest of this work is organized as follows. Section 2 simply gives the iterative least squares algorithm for CMA models. Section 3 derives a matrix decomposition based iterative least squares algorithm. Section 4 provides a numerical example to show the effectiveness of the proposed algorithm. Finally, Section 5 offers some concluding remarks.

2. The iterative least squares algorithm

Consider the following controlled moving average (CMA) system or finite impulse response moving average (FIR-MA) system [18]:

$$y(t) = B(z)u(t) + D(z)v(t), \quad (1)$$

where $\{u(t)\}$ and $\{y(t)\}$ are the input and output sequences of the system, $\{v(t)\}$ is an uncorrelated stochastic noise sequence with zero mean, and $B(z)$ and $D(z)$ are the polynomials, of known orders n_b and n_d , in the unit backward shift operator $z^{-1}[z^{-1}y(t) = y(t-1)]$, defined by

$$B(z) := b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_b} z^{-n_b},$$

$$D(z) := 1 + d_1 z^{-1} + d_2 z^{-2} + \dots + d_{n_d} z^{-n_d}.$$

Assume that n_b and n_d are known and $n := n_b + n_d$, and $y(t) = 0$, $u(t) = 0$ and $v(t) = 0$ for $t \leq 0$. b_i and d_i are the parameters to be estimated from measured input–output data $\{u(t), y(t)\}$.

Define the parameter vector θ and the information vector $\varphi(t)$ as

$$\theta := [b_1, b_2, \dots, b_{n_b}, d_1, d_2, \dots, d_{n_d}]^T \in \mathbb{R}^{n_b+n_d},$$

$$\varphi(t) := [u(t-1), u(t-2), \dots, u(t-n_b), v(t-1), v(t-2), \dots, v(t-n_d)]^T \in \mathbb{R}^{n_b+n_d}.$$

Then Eq. (1) can be written as

$$y(t) = \varphi^T(t)\theta + v(t). \quad (2)$$

Consider a batch of data with length p from $i = t - p + 1$ to $i = t$ and define the stacked output vector $\mathbf{Y}(t)$, the stacked formation matrix $\mathbf{H}(t)$ and the stacked white noise vector $\mathbf{V}(t)$ as

$$\mathbf{Y}(t) := [y(t), y(t-1), \dots, y(t-p+1)]^T \in \mathbb{R}^p,$$

$$\mathbf{H}(t) := [\varphi(t), \varphi(t-1), \dots, \varphi(t-p+1)]^T \in \mathbb{R}^{p \times n},$$

$$\mathbf{V}(t) := [v(t), v(t-1), \dots, v(t-p+1)]^T \in \mathbb{R}^p.$$

From (2), we have

$$\mathbf{Y}(t) = \mathbf{H}(t)\theta + \mathbf{V}(t). \quad (3)$$

Let the norm of the matrix \mathbf{X} be $\|\mathbf{X}\|^2 := \text{tr}[\mathbf{X}\mathbf{X}^T]$ and define a quadratic criterion function,

$$J(\theta) := \|\mathbf{Y}(t) - \mathbf{H}(t)\theta\|^2.$$

Provided that the information vector $\varphi(t)$ is persistently exciting (i.e., $\mathbf{H}^T(t)\mathbf{H}(t)$ is non-singular), minimizing $J(\theta)$ gives the least squares estimate of θ :

$$\hat{\theta}(t) = [\mathbf{H}^T(t)\mathbf{H}(t)]^{-1}\mathbf{H}^T(t)\mathbf{Y}(t). \quad (4)$$

Because $\mathbf{H}(t)$ (i.e., $\varphi(t)$) contains the unknown noise terms $v(t-i)$ in $\varphi(t)$, it is impossible to compute the estimate $\hat{\theta}(t)$ of θ by (4). The solution here is based on the hierarchical identification principle [40,41]. Let $k = 1, 2, 3, \dots$ be an iterative variable, and $\hat{\theta}_k(t)$ the iterative estimate of θ at iteration k ; the unknown $v(t-i)$ in the information vector may be replaced with the estimate $\hat{v}_{k-1}(t-i)$ of $v(t-i)$ at $k-1$, and we define

$$\hat{\varphi}_k(t) := [u(t-1), u(t-2), \dots, u(t-n_b), \hat{v}_{k-1}(t-1), \hat{v}_{k-1}(t-2), \dots, \hat{v}_{k-1}(t-n_d)]^T \in \mathbb{R}^n.$$

From (2), we have $v(t-i) = y(t-i) - \varphi^T(t-i)\theta$. If we use $\hat{\varphi}_k(t-i)$ and $\hat{\theta}_k(t)$ to replace $\varphi(t-i)$ and θ in the above equation, then the estimate $\hat{v}_k(t-i)$ of $v(t-i)$ can be computed using

$$\hat{v}_k(t-i) = y(t-i) - \hat{\varphi}_k^T(t-i)\hat{\theta}_k(t).$$

Define

$$\hat{\mathbf{H}}_k(t) = \begin{bmatrix} \hat{\varphi}_k^T(t) \\ \hat{\varphi}_k^T(t-1) \\ \vdots \\ \hat{\varphi}_k^T(t-p+1) \end{bmatrix} \in \mathbb{R}^{p \times n}.$$

Replacing $\mathbf{H}(t)$ in (4) with $\hat{\mathbf{H}}_k(t)$, we can obtain the following least squares based iterative identification algorithm for CMA (FIR-MA) systems (the CMA-LSI or FIR-MA-LSI algorithm for short):

$$\hat{\boldsymbol{\theta}}_k(t) = [\hat{\mathbf{H}}_k^T(t)\hat{\mathbf{H}}_k(t)]^{-1}\hat{\mathbf{H}}_k^T(t)\mathbf{Y}(t), \quad (5)$$

$$\hat{\mathbf{H}}_k(t) = [\hat{\boldsymbol{\phi}}_k(t), \hat{\boldsymbol{\phi}}_k(t-1), \dots, \hat{\boldsymbol{\phi}}_k(t-p+1)]^T, \quad (6)$$

$$\mathbf{Y}(t) = [y(t), y(t-1), \dots, y(t-p+1)]^T, \quad (7)$$

$$\hat{\boldsymbol{\phi}}_k(t) = [u(t-1), u(t-2), \dots, u(t-n_b), \hat{v}_{k-1}(t-1), \hat{v}_{k-1}(t-2), \dots, \hat{v}_{k-1}(t-n_d)]^T, \quad (8)$$

$$\hat{v}_k(t-i) = y(t-i) - \hat{\boldsymbol{\phi}}_k^T(t-i)\hat{\boldsymbol{\theta}}_k(t). \quad (9)$$

If we let $p = L$ and $t = L$ (L is the length of the data), then we can obtain the CMA-LSI algorithm with data length L :

$$\hat{\boldsymbol{\theta}}_k = [\hat{\mathbf{H}}_k^T\hat{\mathbf{H}}_k]^{-1}\hat{\mathbf{H}}_k^T\mathbf{Y}, \quad (10)$$

$$\hat{\mathbf{H}}_k = [\hat{\boldsymbol{\phi}}_k(L), \hat{\boldsymbol{\phi}}_k(L-1), \dots, \hat{\boldsymbol{\phi}}_k(1)]^T, \quad (11)$$

$$\mathbf{Y} = [y(L), y(L-1), \dots, y(1)]^T, \quad (12)$$

$$\hat{\boldsymbol{\phi}}_k(t) = [u(t-1), u(t-2), \dots, u(t-n_b), \hat{v}_{k-1}(t-1), \hat{v}_{k-1}(t-2), \dots, \hat{v}_{k-1}(t-n_d)]^T, \quad (13)$$

$$\hat{v}_k(t) = y(t) - \hat{\boldsymbol{\phi}}_k^T(t)\hat{\boldsymbol{\theta}}_k, \quad t = 1, 2, \dots, L. \quad (14)$$

The algorithm in (10)–(14) requires computing the inverse of large matrices $[\hat{\mathbf{H}}_k^T\hat{\mathbf{H}}_k]$ with sizes $(n_b + n_d) \times (n_b + n_d)$ at each iteration $k = 1, 2, \dots$. We note that the $(1, 1)$ th partitioned matrix in $\hat{\mathbf{H}}_k^T\hat{\mathbf{H}}_k$ does not depend on iteration k —see the matrix $\mathbf{S} := \Phi^T\Phi$ in Eq. (15). The following derives the matrix decomposition based iterative least squares algorithm which only requires computing the inverse of an $n_d \times n_d$ matrix.

3. The matrix decomposition based iterative least squares algorithm

Define the input information vector $\boldsymbol{\phi}(t)$ and the noise vector $\boldsymbol{\psi}_k(t)$ at iteration k as

$$\begin{aligned} \boldsymbol{\phi}(t) &:= [u(t-1), u(t-2), \dots, u(t-n_b)]^T \in \mathbb{R}^{n_b}, \\ \hat{\boldsymbol{\psi}}_k(t) &:= [\hat{v}_{k-1}(t-1), \hat{v}_{k-1}(t-2), \dots, \hat{v}_{k-1}(t-n_d)]^T \in \mathbb{R}^{n_d}, \end{aligned}$$

and the stacked information matrices as

$$\begin{aligned} \Phi &:= [\boldsymbol{\phi}(L), \boldsymbol{\phi}(L-1), \dots, \boldsymbol{\phi}(1)]^T \in \mathbb{R}^{L \times n_b}, \\ \hat{\Psi}_k &:= [\hat{\boldsymbol{\psi}}_k(L), \hat{\boldsymbol{\psi}}_k(L-1), \dots, \hat{\boldsymbol{\psi}}_k(1)]^T \in \mathbb{R}^{L \times n_d}. \end{aligned}$$

Then $\hat{\boldsymbol{\phi}}_k(t)$ in (13) and $\hat{\mathbf{H}}_k(t)$ in (11) can be expressed as

$$\begin{aligned} \hat{\boldsymbol{\phi}}_k(t) &= \begin{bmatrix} \boldsymbol{\phi}(t) \\ \hat{\boldsymbol{\psi}}_k(t) \end{bmatrix} = [\boldsymbol{\phi}^T(t), \hat{\boldsymbol{\psi}}_k^T(t)]^T \in \mathbb{R}^{n_b+n_d}, \\ \hat{\mathbf{H}}_k &= [\hat{\boldsymbol{\phi}}_k(L), \hat{\boldsymbol{\phi}}_k(L-1), \dots, \hat{\boldsymbol{\phi}}_k(1)]^T \\ &= \begin{bmatrix} \boldsymbol{\phi}(L) & \boldsymbol{\phi}(L-1) & \cdots & \boldsymbol{\phi}(1) \\ \hat{\boldsymbol{\psi}}_k(L) & \hat{\boldsymbol{\psi}}_k(L-1) & \cdots & \hat{\boldsymbol{\psi}}_k(1) \end{bmatrix}^T \\ &= \begin{bmatrix} \boldsymbol{\phi}^T(L) & \hat{\boldsymbol{\psi}}_k^T(L) \\ \boldsymbol{\phi}^T(L-1) & \hat{\boldsymbol{\psi}}_k^T(L-1) \\ \vdots & \vdots \\ \boldsymbol{\phi}^T(1) & \hat{\boldsymbol{\psi}}_k^T(1) \end{bmatrix} = [\Phi, \hat{\Psi}_k] \in \mathbb{R}^{L \times n}. \end{aligned}$$

Define the data product moment matrix

$$\begin{aligned} \mathbf{S}_k &:= \hat{\mathbf{H}}_k^T\hat{\mathbf{H}}_k = \begin{bmatrix} \Phi^T \\ \hat{\Psi}_k^T \end{bmatrix} [\Phi, \hat{\Psi}_k] \\ &= \begin{bmatrix} \Phi^T\Phi & \Phi^T\hat{\Psi}_k \\ \hat{\Psi}_k^T\Phi & \hat{\Psi}_k^T\hat{\Psi}_k \end{bmatrix} = \begin{bmatrix} \mathbf{S} & \Phi^T\hat{\Psi}_k \\ \hat{\Psi}_k^T\Phi & \hat{\Psi}_k^T\hat{\Psi}_k \end{bmatrix} \in \mathbb{R}^{(n_b+n_d) \times (n_b+n_d)}, \end{aligned} \quad (15)$$

where $\mathbf{S} := \Phi^T \Phi$. Eq. (10) can be expressed as

$$\hat{\theta}_k = [\hat{\mathbf{H}}_k^T \hat{\mathbf{H}}_k]^{-1} \hat{\mathbf{H}}_k^T \mathbf{Y} = \mathbf{S}_k^{-1} \begin{bmatrix} \Phi^T \\ \hat{\Psi}_k^T \end{bmatrix} \mathbf{Y} = \mathbf{S}_k^{-1} \begin{bmatrix} \Phi^T \mathbf{Y} \\ \hat{\Psi}_k^T \mathbf{Y} \end{bmatrix}. \quad (16)$$

When we use the algorithm in (15)–(16) to compute the parameter estimation vector $\hat{\theta}_k$ with iteration k increasing, it is necessary to compute the inverse of the data product moment matrix \mathbf{S}_k at each iteration. This will lead to a heavy computational burden because the (1, 1)th partitioned matrix $\mathbf{S} := \Phi^T \Phi$ of \mathbf{S}_k is constant for each k (i.e., independent of k). This motivates us to study the matrix decomposition based estimation algorithm so as to reduce the computational burden of the algorithm in (15)–(16).

Lemma 1. Suppose that $\mathbf{A}_1 \in \mathbb{R}^{m \times m}$ and $\mathbf{Q} := \mathbf{A}_2 - \mathbf{A}_{21} \mathbf{A}_1^{-1} \mathbf{A}_{12} \in \mathbb{R}^{n \times n}$ are nonsingular matrices. Then the following block matrix inversion relation holds:

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_2 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_1^{-1} + \mathbf{A}_1^{-1} \mathbf{A}_{12} \mathbf{Q}^{-1} \mathbf{A}_{21} \mathbf{A}_1^{-1} & -\mathbf{A}_1^{-1} \mathbf{A}_{12} \mathbf{Q}^{-1} \\ -\mathbf{Q}^{-1} \mathbf{A}_{21} \mathbf{A}_1^{-1} & \mathbf{Q}^{-1} \end{bmatrix}.$$

Applying Lemma 1 to (15) gives

$$\mathbf{S}_k^{-1} = \begin{bmatrix} \mathbf{S} & \Phi^T \hat{\Psi}_k \\ \hat{\Psi}_k^T \Phi & \hat{\Psi}_k^T \hat{\Psi}_k \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{S}^{-1} + \mathbf{S}^{-1} \Phi^T \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \Phi \mathbf{S}^{-1} & -\mathbf{S}^{-1} \Phi^T \hat{\Psi}_k \mathbf{Q}_k^{-1} \\ -\mathbf{Q}_k^{-1} \hat{\Psi}_k^T \Phi \mathbf{S}^{-1} & \mathbf{Q}_k^{-1} \end{bmatrix},$$

$$\mathbf{Q}_k := \hat{\Psi}_k^T \hat{\Psi}_k - \hat{\Psi}_k^T \Phi \mathbf{S}^{-1} \Phi^T \hat{\Psi}_k = \hat{\Psi}_k^T [\mathbf{I} - \Phi \mathbf{S}^{-1} \Phi^T] \hat{\Psi}_k,$$

where \mathbf{I} is an identity matrix of appropriate size.

Define the constant vectors and matrices: $\alpha := \mathbf{S}^{-1} \Phi^T \mathbf{Y} \in \mathbb{R}^{n_b}$, $\beta := \Phi \alpha \in \mathbb{R}^{L \times n_b}$, $\mathbf{R} := \mathbf{S}^{-1} \Phi^T \in \mathbb{R}^{n_b \times L}$, $\mathbf{M} := \mathbf{I} - \Phi \mathbf{S}^{-1} \Phi^T \in \mathbb{R}^{L \times L}$, which are independent of k and are computed once at the first iteration. From (16), we have

$$\begin{aligned} \hat{\theta}_k &= \begin{bmatrix} \mathbf{S}^{-1} + \mathbf{S}^{-1} \Phi^T \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \Phi \mathbf{S}^{-1} & -\mathbf{S}^{-1} \Phi^T \hat{\Psi}_k \mathbf{Q}_k^{-1} \\ -\mathbf{Q}_k^{-1} \hat{\Psi}_k^T \Phi \mathbf{S}^{-1} & \mathbf{Q}_k^{-1} \end{bmatrix} \begin{bmatrix} \Phi^T \mathbf{Y} \\ \hat{\Psi}_k^T \mathbf{Y} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{S}^{-1} \Phi^T \mathbf{Y} + \mathbf{S}^{-1} \Phi^T \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \Phi \mathbf{S}^{-1} \Phi^T \mathbf{Y} - \mathbf{S}^{-1} \Phi^T \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \mathbf{Y} \\ -\mathbf{Q}_k^{-1} \hat{\Psi}_k^T \Phi \mathbf{S}^{-1} \Phi^T \mathbf{Y} + \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \mathbf{Y} \end{bmatrix} \\ &= \begin{bmatrix} \alpha + \mathbf{S}^{-1} \Phi^T \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \Phi \alpha - \mathbf{S}^{-1} \Phi^T \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \mathbf{Y} \\ -\mathbf{Q}_k^{-1} \hat{\Psi}_k^T \Phi \alpha + \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \mathbf{Y} \end{bmatrix} \\ &= \begin{bmatrix} \alpha + \mathbf{R} \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \beta - \mathbf{R} \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \mathbf{Y} \\ -\mathbf{Q}_k^{-1} \hat{\Psi}_k^T \beta + \mathbf{Q}_k^{-1} \hat{\Psi}_k^T \mathbf{Y} \end{bmatrix} \\ &= \begin{bmatrix} \alpha + \mathbf{R} \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T (\beta - \mathbf{Y}) \\ -\mathbf{Q}_k^{-1} \hat{\Psi}_k^T (\beta - \mathbf{Y}) \end{bmatrix}. \end{aligned}$$

Thus, we can summarize the matrix decomposition based iterative least squares (MD-LSI) algorithm:

$$\hat{\theta}_k = \begin{bmatrix} \alpha + \mathbf{R} \hat{\Psi}_k \mathbf{Q}_k^{-1} \hat{\Psi}_k^T (\beta - \mathbf{Y}) \\ -\mathbf{Q}_k^{-1} \hat{\Psi}_k^T (\beta - \mathbf{Y}) \end{bmatrix}, \quad k = 1, 2, 3, \dots, \quad (17)$$

$$\hat{\Psi}_k := [\hat{\psi}_k(L), \hat{\psi}_k(L-1), \dots, \hat{\psi}_k(1)]^T, \quad (18)$$

$$\mathbf{Q}_k = \hat{\Psi}_k^T \mathbf{M} \hat{\Psi}_k, \quad (19)$$

$$\phi(t) = [u(t-1), u(t-2), \dots, u(t-n_b)]^T, \quad t = 1, 2, \dots, L, \quad (20)$$

$$\hat{\psi}_k(t) = [\hat{v}_{k-1}(t-1), \hat{v}_{k-1}(t-2), \dots, \hat{v}_{k-1}(t-n_d)]^T, \quad (21)$$

$$\hat{\phi}_k(t) = \begin{bmatrix} \phi(t) \\ \hat{\psi}_k(t) \end{bmatrix} = [\phi^T(t), \hat{\psi}_k^T(t)]^T \in \mathbb{R}^{n_b+n_d}, \quad (22)$$

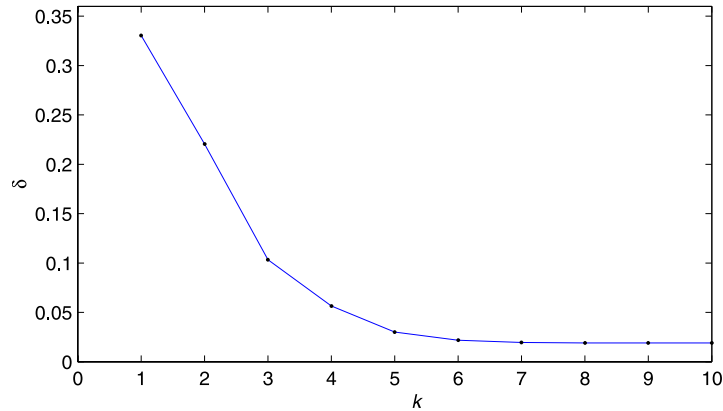
$$\hat{v}_k(t) = y(t) - \hat{\phi}_k^T(t) \hat{\theta}_k, \quad (23)$$

$$\mathbf{Y} = [y(L), y(L-1), \dots, y(1)]^T, \quad (24)$$

$$\Phi = [\phi(L), \phi(L-1), \dots, \phi(1)]^T, \quad (25)$$

Table 1The MD-LSI parameter estimates and errors with $L = 1000$.

k	b_1	b_2	d_1	d_2	δ (%)
1	1.69794	2.31247	−0.00393	0.00421	33.04023
2	1.70971	2.31118	−0.46880	0.13338	22.04944
3	1.70870	2.30967	−0.47665	0.51366	10.33702
4	1.69559	2.31189	−0.56231	0.62782	5.65591
5	1.69188	2.31794	−0.58127	0.71091	3.01200
6	1.68972	2.31902	−0.59098	0.73631	2.18613
7	1.68937	2.31989	−0.59286	0.74514	1.95503
8	1.68929	2.31994	−0.59333	0.74690	1.90875
9	1.68927	2.31996	−0.59341	0.74730	1.89897
10	1.68927	2.31996	−0.59343	0.74738	1.89708
True values	1.68000	2.32000	−0.64000	0.78000	

**Fig. 1.** The MD-LSI estimation error δ versus k with $L = 1000$.

$$\mathbf{S} = \Phi^T \Phi, \quad \alpha = \mathbf{S}^{-1} \Phi^T \mathbf{Y}, \quad \beta = \Phi \alpha, \quad (26)$$

$$\mathbf{R} = \mathbf{S}^{-1} \Phi^T, \quad \mathbf{M} = \mathbf{I} - \Phi \mathbf{S}^{-1} \Phi^T. \quad (27)$$

The steps of computing the parameter estimate $\hat{\theta}_k$ in the MD-LSI algorithm are listed in the following.

1. Collect the input–output data $\{u(t), y(t): t = 1, 2, 3, \dots, L\}$.
2. To initialize: let $k = 1$, $\hat{v}_0(t)$ be a random number, and give a small positive number ε .
3. Form \mathbf{Y} by using (24), $\phi(t)$ by using (20) and Φ by using (25).
4. Compute \mathbf{S} , α , β , \mathbf{R} and \mathbf{M} by using (26)–(27).
5. Form $\hat{\psi}_k(t)$ by using (21) and $\hat{\Psi}_k$ by using (18), and compute \mathbf{Q}_k by using (19).
6. Update the parameter estimate $\hat{\theta}_k$ by using (17) and compute $\hat{v}_k(t)$ by using (23).
7. If $\|\hat{\theta}_k - \hat{\theta}_{k-1}\| \leq \varepsilon$, terminate this procedure and obtain the iterative variable k and the parameter estimate $\hat{\theta}_k$; otherwise increase k by 1 and return to step 5.

4. A numerical example

Consider the following CMA system:

$$y(t) = B(z)u(t) + D(z)v(t),$$

$$B(z) = b_1 z^{-1} + b_2 z^{-2} = 1.680z^{-1} + 2.32z^{-2},$$

$$D(z) = 1 + d_1 z^{-1} + d_2 z^{-2} = 1 - 0.64z^{-1} + 0.78z^{-2},$$

$$\theta = [b_1, b_2, d_1, d_2]^T = [1.68, 2.32, -0.64, 0.78]^T.$$

The input $\{u(t)\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance, and $\{v(t)\}$ as a white noise sequence with zero mean and variance $\sigma^2 = 0.50$ and the noise-to-signal ratio of the system is $\delta_{ns} = 24.80\%$. Applying the proposed MD-LSI algorithm to estimate the parameters of this CMA system, the parameter estimates and their estimation errors are shown in Table 1 with the data length $L = 1000$, and the estimation error $\delta := \|\hat{\theta}_k - \theta\| / \|\theta\|$ versus k is shown in Fig. 1.

From Table 1 and Fig. 1, we can draw the following conclusion: as the iteration k increases, the estimation error becomes smaller. This shows that the proposed iterative least squares algorithm works well for estimating the parameters of controlled moving average systems.

5. Conclusions

This work uses the partitioned matrix inversion formula to derive an iterative least squares parameter estimation algorithm for controlled moving average systems based on matrix decomposition. The proposed algorithm has a high computational efficiency and can be used to study identification problems for other systems [42–49].

References

- [1] J.H. Li, R.F. Ding, Y. Yang, Iterative parameter identification methods for nonlinear functions, *Applied Mathematical Modelling* 26 (6) (2012) 2739–2750.
- [2] Y.J. Liu, Y.S. Xiao, X.L. Zhao, Multi-innovation stochastic gradient algorithm for multiple-input single-output systems using the auxiliary model, *Applied Mathematics and Computation* 215 (4) (2009) 1477–1483.
- [3] Y.J. Liu, J. Sheng, R.F. Ding, Convergence of stochastic gradient estimation algorithm for multivariable ARX-like systems, *Computers & Mathematics with Applications* 59 (8) (2010) 2615–2627.
- [4] J. Ding, L.L. Han, X.M. Chen, Time series AR modeling with missing observations based on the polynomial transformation, *Mathematical and Computer Modelling* 51 (5–6) (2010) 527–536.
- [5] F. Ding, T. Chen, Hierarchical identification of lifted state-space models for general dual-rate systems, *IEEE Transactions on Circuits and Systems—I: Regular Papers* 52 (6) (2005) 1179–1187.
- [6] Y. Shi, H. Fang, Kalman filter based identification for systems with randomly missing measurements in a network environment, *International Journal of Control* 83 (3) (2010) 538–551.
- [7] B. Yu, Y. Shi, H. Huang, l -2 and l -infinity filtering for multirate systems using lifted models, *Circuits, Systems, and Signal Processing* 27 (5) (2008) 699–711.
- [8] Y. Shi, B. Yu, Output feedback stabilization of networked control systems with random delays modeled by Markov chains, *IEEE Transactions on Automatic Control* 54 (7) (2009) 1668–1674.
- [9] M. Yan, Y. Shi, Robust discrete-time sliding mode control for uncertain systems with time-varying state delay, *IET Control Theory & Applications* 2 (8) (2008) 662–674.
- [10] H. Fang, J. Wu, Y. Shi, Genetic adaptive state estimation with missing input/output data, *Proceedings of the Institution of Mechanical Engineers Part I: Journal of Systems and Control Engineering* 224 (5) (2010) 611–617.
- [11] J.B. Zhang, F. Ding, Y. Shi, Self-tuning control based on multi-innovation stochastic gradient parameter estimation, *Systems & Control Letters* 58 (1) (2009) 69–75.
- [12] F. Ding, T. Chen, Z. Iwai, Adaptive digital control of Hammerstein nonlinear systems with limited output sampling, *SIAM Journal on Control and Optimization* 45 (6) (2007) 2257–2276.
- [13] H.H. Yin, Z.F. Zhu, et al., Model order determination using the Hankel matrix of impulse responses, *Applied Mathematics Letters* 24 (5) (2011) 797–802.
- [14] F. Ding, X.P. Liu, G. Liu, Identification methods for Hammerstein nonlinear systems, *Digital Signal Processing* 21 (2) (2011) 215–238.
- [15] Y.S. Xiao, Y. Zhang, J. Ding, J.Y. Dai, The residual based interactive least squares algorithms and simulation studies, *Computers & Mathematics with Applications* 58 (6) (2009) 1190–1197.
- [16] D.Q. Wang, Least squares-based recursive and iterative estimation for output error moving average (OEMA) systems using data filtering, *IET Control Theory and Applications* 5 (14) (2011) 1648–1657.
- [17] F. Ding, H.Z. Yang, F. Liu, Performance analysis of stochastic gradient algorithms under weak conditions, *Science in China Series F—Information Sciences* 51 (9) (2008) 1269–1280.
- [18] L.Y. Wang, L. Xie, X.F. Wang, The residual based interactive stochastic gradient algorithms for controlled moving average models, *Applied Mathematics and Computation* 211 (2) (2009) 442–449.
- [19] F. Ding, G. Liu, X.P. Liu, Partially coupled stochastic gradient identification methods for non-uniformly sampled systems, *IEEE Transactions on Automatic Control* 55 (8) (2010) 1976–1981.
- [20] F. Ding, T. Chen, Performance analysis of multi-innovation gradient type identification methods, *Automatica* 43 (1) (2007) 1–14.
- [21] F. Ding, X.P. Liu, G. Liu, Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises, *Signal Processing* 89 (10) (2009) 1883–1890.
- [22] F. Ding, G. Liu, X.P. Liu, Parameter estimation with scarce measurements, *Automatica* 47 (8) (2011) 1646–1655.
- [23] F. Ding, Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling, *Applied Mathematical Modelling* (2012) <http://dx.doi.org/10.1016/j.apm.2012.04.039>.
- [24] D.Q. Wang, G.W. Yang, R.F. Ding, Gradient-based iterative parameter estimation for Box–Jenkins systems, *Computers & Mathematics with Applications* 60 (5) (2010) 1200–1208.
- [25] F. Ding, Y.J. Liu, B. Bao, Gradient based and least squares based iterative estimation algorithms for multi-input multi-output systems, *Proceedings of the Institution of Mechanical Engineers Part I: Journal of Systems and Control Engineering* 226 (1) (2012) 43–55.
- [26] D.Q. Wang, F. Ding, Least squares based and gradient based iterative identification for Wiener nonlinear systems, *Signal Processing* 91 (5) (2011) 1182–1189.
- [27] M. Dehghan, M. Hajarian, Two algorithms for finding the Hermitian reflexive and skew-Hermitian solutions of Sylvester matrix equations, *Applied Mathematics Letters* 24 (4) (2011) 444–449.
- [28] M. Dehghan, M. Hajarian, An iterative method for solving the generalized coupled Sylvester matrix equations over generalized bisymmetric matrices, *Applied Mathematical Modelling* 34 (3) (2010) 639–654.
- [29] G.H. Golub, C.F. Van Loan, *Matrix Computations*, third ed., Johns Hopkins University Press, Baltimore, MD, 1996.
- [30] F. Ding, X.P. Liu, J. Ding, Iterative solutions of the generalized Sylvester matrix equations by using the hierarchical identification principle, *Applied Mathematics and Computation* 197 (1) (2008) 41–50.
- [31] F. Ding, P.X. Liu, G. Liu, Gradient based and least-squares based iterative identification methods for OE and OEMA systems, *Digital Signal Processing* 20 (3) (2010) 664–677.
- [32] Y.J. Liu, D.Q. Wang, et al., Least-squares based iterative algorithms for identifying Box–Jenkins models with finite measurement data, *Digital Signal Processing* 20 (5) (2010) 1458–1467.
- [33] B. Bao, Y.Q. Xu, J. Sheng, R.F. Ding, Least squares based iterative parameter estimation algorithm for multivariable controlled ARMA system modelling with finite measurement data, *Mathematical and Computer Modelling* 53 (9–10) (2011) 1664–1669.
- [34] L. Xie, Y.J. Liu, H.Z. Yang, Gradient based and least squares based iterative algorithms for matrix equations $AXB + CX^TD = F$, *Applied Mathematics and Computation* 217 (5) (2010) 2191–2199.
- [35] F. Ding, Transformations between some special matrices, *Computers & Mathematics with Applications* 59 (8) (2010) 2676–2695.
- [36] Z.N. Zhang, F. Ding, X.G. Liu, Hierarchical gradient based iterative parameter estimation algorithm for multivariable output error moving average systems, *Computers & Mathematics with Applications* 61 (3) (2011) 672–682.
- [37] J. Ding, F. Ding, X.P. Liu, G. Liu, Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data, *IEEE Transactions on Automatic Control* 56 (11) (2011) 2677–2683.
- [38] J. Ding, F. Ding, Bias compensation based parameter estimation for output error moving average systems, *International Journal of Adaptive Control and Signal Processing* 25 (12) (2011) 1100–1111.

- [39] H.Y. Hu, R.F. Ding, Decomposition based iterative estimation algorithm for autoregressive moving average models, in: Proceedings of the 31st Chinese Control Conference, CCC2012, Hefei, China, July 25–27, 2012.
- [40] F. Ding, T. Chen, Hierarchical gradient-based identification of multivariable discrete-time systems, *Automatica* 41 (2) (2005) 315–325.
- [41] F. Ding, T. Chen, Hierarchical least squares identification methods for multivariable systems, *IEEE Transactions on Automatic Control* 50 (3) (2005) 397–402.
- [42] W. Wang, F. Ding, J.Y. Dai, Maximum likelihood least squares identification for systems with autoregressive moving average noise, *Applied Mathematical Modelling* 36 (5) (2012) 1842–1853.
- [43] J.X. Ma, F. Ding, Recursive relations of the cost functions for the least squares algorithms for multivariable systems, *Circuits, Systems and Signal Processing* (2012) <http://dx.doi.org/10.1007/s00034-012-9448-4>.
- [44] F. Ding, X.P. Liu, Y. Shi, Convergence analysis of estimation algorithms for dual-rate stochastic systems, *Applied Mathematics and Computation* 176 (1) (2006) 245–261.
- [45] Y.J. Liu, F. Ding, Y. Shi, Least squares estimation for a class of non-uniformly sampled systems based on the hierarchical identification principle, *Circuits, Systems and Signal Processing* (2012) <http://dx.doi.org/10.1007/s00034-012-9421-2>.
- [46] F. Ding, Y. Shi, T. Chen, Performance analysis of estimation algorithms of non-stationary ARMA processes, *IEEE Transactions on Signal Processing* 54 (3) (2006) 1041–1053.
- [47] J.H. Li, F. Ding, Maximum likelihood stochastic gradient estimation for Hammerstein systems with colored noise based on the key term separation technique, *Computers & Mathematics with Applications* 62 (11) (2011) 4170–4177.
- [48] F. Ding, Y. Gu, Performance analysis of the auxiliary model based least squares identification algorithm for one-step state delay systems, *International Journal of Computer Mathematics* (2012) <http://dx.doi.org/10.1080/00207160.2012.698008>.
- [49] J.H. Li, F. Ding, G.W. Yang, Maximum likelihood least squares identification method for input nonlinear finite impulse response moving average systems, *Mathematical and Computer Modelling* 55 (3–4) (2012) 442–450.